Analyse et traitement vidéo : Mouvement Humain et Relevé de Températures

-

Rapport de stage

Jacomme Florian

31 août 2011

Table des matières

I Le projet **6**

1 Le cadre **7**
1.1 Le projet Pré-Therm 7
1.2 Le stage 8

2 Les composantes **9**
2.1 Le corpus de vidéos 9
2.2 La caméra infrarouge 10
2.3 Les outils 10

II Les pré-traitements **11**

3 La détection du couffin **12**
3.1 Problématique 12
3.2 La détection de lignes 12
3.2.1 Les contours 12
3.2.2 La transformée de Hough 13

4 Redressement et découpage **14**
4.1 Les coordonnées spatiales 14
4.2 Correction de la perspective 16
4.2.1 La déformation 16
4.2.2 Implémentation 16

5 La détection du nouveau-né **17**
5.1 La silhouette 17
5.1.1 Les vidéos du corpus 17
5.1.2 Les images thermiques 17
5.2 Squelettisation 18

III La position et le mouvement **19**

6 Modelisation du corps humain **20**
6.1 Les modèles 20
6.1.1 Modèles courant 20
6.1.2 Modèle adopté 21
6.2 Les contraintes 22

7 Détection de posture par squelette actif **23**
7.1 Principe de fonctionnement 23
7.1.1 Snake 23
7.1.2 Adaptation au problème 24
7.2 Implémentation 24
7.2.1 La minimisation de l'énergie 24
7.2.2 Problèmes 25

8 Initialisation et précision **26**
8.1 Semi-supervisée 26
8.2 Le flot optique 26
8.3 Utilisation de l'estimation 27
8.3.1 Initialisation 27
8.3.2 Optimisation de la convergence 27

IV Applications **28**

9 Le mouvement dans les vidéos du corpus **29**
9.1 Les contraintes et les problèmes 29
9.1.1 But du corpus 29
9.1.2 Précision et fiabilité de la détection 29
9.2 Calcul du mouvement et extension à la classification 30
9.2.1 Mouvement global 30
9.2.2 Découpage en zone 30

10 Position et température **31**
10.1 Les vidéos thermiques 31
10.2 Températures 32
10.2.1 Relevés globaux 32
10.2.2 Relevés locaux 32
10.3 Utilisation des relevés 33

V Annexes **36**

A Les difficultés des vidéos du corpus **37**

B Les images thermiques **40**
B.1 Les bases de la thermographie infrarouge 40
B.2 La caméra Optris PI 40

C Le progamme développé pour les tests **42**
C.1 Transformée de Hough 42
C.2 Résolution du système linéaire pour le recadrage 42
C.3 Algorithme squelette actif 43

D Exemples de résultats **44**

Introduction

L'acquisition du master 2 nécessite de valider les connaissances reçues pendant l'année de formation. Cette validation s'effectue par le biais d'un stage dans le domaine de la recherche en imagerie numérique. L'étudiant peut ainsi parfaire ses compétences et bien souvent en acquérir de nouvelles. Outre l'aspect académique, c'est aussi l'occasion de découvrir le monde du travail et de la recherche, tout en étant guidé et encadré par une personne expérimentée.

L'université de Caen

L'Université de Caen est un établissement multidisciplinaire regroupant près de 25 000 étudiants, répartis sur cinq sites dans l'agglomération caennaise et dans plusieurs villes de la région. Le département informatique, au sein duquel a été effectué ce stage, se situe au Campus 2, le pôle Sciences et Technologies. En plus d'accueillir les salles de cours et de TP, de nombreux laboratoires et bureaux s'y trouvent.

Les intervenants du projet

Trois personnes sont impliquées dans ce stage :

M. Chahir Youssef : maître de stage. Il est professeur au département informatique de l'université et est également responsable du Master 2 Informatique Rercherche "Langage, Image et Document" dans le cadre duquel s'est déroulé ce stage. Il occupe un poste de chercheur au sein du GREYC [1], laboratoire associé au CNRS [2] travaillant sur de nombreux domaines allant de l'algorithmique à l'instrumentation, en passant bien évidemment par le traitement et l'analyse d'images.

M. Jouen François est directeur d'études à l'EPHE [3], Section des Sciences de la Vie et de la Terre et directeur scientifique chez Grey-Soft. Il est spé-

1. Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen
2. Centre national de la recherche scientifique
3. École Pratique des Hautes Études

cialisé dans la psychologie du développement, et supervise de nombreux projets de recherche, sur le développement des bébés prématurés et des nouveaux-nés immunodéficients. La société Grey-Soft est située à Paris. Ils conçoivent, développent et fabriquent des solutions de systèmes embarqués, notamment dans le domaine des instruments d'imagerie médicale.

M. Jacomme Florian : Après avoir obtenu successivement un DUT Informatique et une licence professionnelle en Image Numérique, je suis actuellement étudiant en Master 2 Informatique RLID. Ce stage est pour moi l'occasion de m'immerger dans le monde de la recherche, jusqu'alors inconnu. Il me permet par ailleurs de travailler sur mon domaine de prédilection : l'imagerie numérique.

L'imagerie médicale

Parmi les nombreux domaines d'application de l'informatique, l'image occupe une place de plus en plus importante. Que ce soient les secteurs de la médecine, l'archéologie, l'aéronautique, la photographie, les médias... tous ont à présent recours à l'image numérique pour analyser, traiter, visualiser les informations, voire pour créer de nouveaux contenus.

Voir un problème permet souvent de mieux le diagnostiquer, et peut aussi aider à le résoudre. L'imagerie médicale a pour but de créer une représentation visuelle d'une information à caratère médical, la rendant ainsi interprétable. La médecine suscite et profite de nombreuses avancées dans la recherche en analyse et traitement des images. Que ce soit pour voir les volumes de l'imagerie par résonnance magnétique ou pour détecter les cellules cancéreuses sur une image de microscope, les applications ne manquent pas.

Première partie

Le projet

Chapitre 1

Le cadre

1.1 Le projet Pré-Therm

Chez les nouveaux-nés, et plus spécifiquement les prématurés, de nombreux indicateurs de l'état de santé peuvent être exploités pour aider au développement, tels que le rythme cardiaque, la fréquence respiratoire, la saturation en oxygène. À ceux-ci peuvent s'ajouter deux autres mesures, auquelles s'intéresse ce projet : les mouvements ainsi que les variations de température sur le corps. Leurs caractéristiques peuvent en effet être utilisées afin de diagnostiquer certaines maladies [1]. De plus, leur combinaison permet de déduire l'état de stress et le niveau de douleur, information précieuse dans le cas des nourrisons, qui ne disposent pas de beaucoup de moyen de l'exprimer.

Le projet Mesures Thermiques de la Douleur et du Stress chez le Prématuré (Pré-Therm) est un projet de recherche sur 3 ans lancé en 2009. Il s'organise autour de deux axes de recherche :

- La mise au point d'un système de capteur thermique adapté à l'environnement hospitalier et à la mesure thermique chez le bébé prématuré. La construction de ce système nécessite le développement d'algorithmes nouveaux de traitement des images infrarouges.
- La mise au point de l'intégration des signaux thermiques, physiologiques et comportementaux dans le but d'obtenir un indicateur fiable du stress et de la douleur chez le bébé prématuré.

Quatre partenaires différents collaborent sur ce projet : le laboratoire PALM (Psychologie des Actions Langagières et Motrices) de l'Université de Caen, le GREYC, le CHU de Caen et Grey-Soft. Ils se répartissent les tâches de la manière suivante :

PALM : coordination du projet, recueil des données expérimentales

CHU : recrutement des sujets et évalation clinique

Grey-Soft : construction du système capteur thermique

GREYC : analyse des images infrarouges, intégration des signaux physiologique et comportementaux.

C'est dans le cadre de la tâche incombant au GREYC que se situe ce stage.

1.2 Le stage

Face à la fragilité inhérente à ce type de patient, les méthodes de mesures non-invasives sont fortement privilégiées. La thermographie infrarouge (cf. annexe B) permet d'obtenir une vidéo renseignant à la fois les mouvements du nourrisson et les températures à la surface de sa peau, sans affecter le sujet.

Les images infrarouges n'ayant pas encore été réalisées, le but de ce stage est avant tout théorique et relève de l'étude de faisabilité et des solutions utilisables. Deux grandes lignes de recherche se dégagent : étudier les possiblités de détection du mouvement et leurs liens avec le diagnostique, et faire correspondre les relevés de température avec leurs emplacements sur le corps et leurs éventuelles significations.

Le travail porte donc en premier lieu sur la segmentation de l'image, afin de détecter le nouveau-né. Nous cherchons ensuite à estimer sa posture, c'est-à-dire la position de chaque membre. Cela nous permet alors de suivre ses mouvements mais aussi de connaître les variations de températures locales, comme par exemple une baisse de chaleur dans un bras ou une jambe.

Chapitre 2

Les composantes

2.1 Le corpus de vidéos

Afin d'étudier le mouvement biologique des nouveaux-nés, un corpus de vidéos a été mis à disposition. Il s'agit de séquences montrant chacune un bébé dans un couffin d'hôpital, durant cinq à quinze minutes. La scène est filmée d'un point fixe en plongée. Les vidéos ne sont pas triées et rien, si ce n'est le diagnostique d'un expert après visualisation, ne permet de savoir si nous avons affaire à un individu sain ou affecté.

Les films étaient stockés bout à bout sur des DVD. Il a donc fallut, en tant que travail préliminaire, les transférer sur un autre support et les découper afin d'avoir une vidéo par enfant.

Suite à ce premier visionnage, on peut constater que la prise de vue ne suit pas de protocole défini : l'angle de vue, les vêtements, le couffin, l'âge du bébé... tout peut varier d'une séquence à l'autre.

FIGURE 2.1 – Exemples de vidéos du corpus

2.2 La caméra infrarouge

La caméra utilisée est une *Optris PI*[1]. Elle peut détecter avec les réglages par défaut les températures de -20 ° C à 100 ° C avec une précision de 2% et une fréquence de 120Hz. Ces paramètres sont adaptables en fonction des besoins. Les images en sorties ont une résolution de 160x120 pixels. Plusieurs optiques différentes peuvent être utilisées afin de modififer le champs de vision et la distance focale(cf. annexe). Elle fonctionne en 5V via une interface USB 2.0.

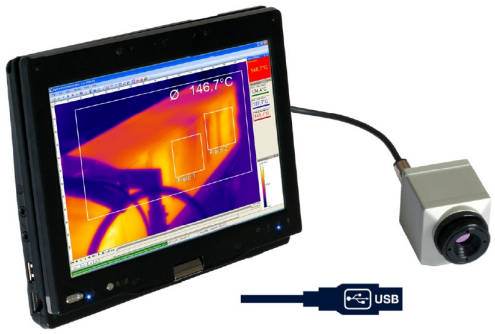

FIGURE 2.2 – La caméra thermique OPTRIS PI

2.3 Les outils

Afin de réaliser les différents tests et développements nécessaires à ce projet, plusieurs choix techniques ont été faits. Tout d'abord, le choix du langage de programmation. Le traitement d'une vidéo est un processus lourd et coûteux en ressources. La plupart des développements seront réalisés dans un but de test et la recherche permet donc rarement d'avoir directement des algorithmes optimisés. Le C++ a été choisi car il est relativement rapide et puissant, et largement supporté.

Le but du projet n'étant pas de réinventer des techniques de traitements d'images déjà existantes, des bibliothèques permettant d'effectuer les manipulations de bases ont été utilisées. Les deux principaux outils choisis sont Pandore [2] et CImg [3]. Tous deux développés au sein du GREYC, ils permettent de manipuler les images de toutes sortes, et implémentent la plupart des opérations de base, tels que le filtrage, la transformation, la segmentation... Ils disposent tout deux d'une abondante documentation, et sont utilisables en C++.

1. http ://www.optris.fr
2. http ://www.greyc.ensicaen.fr/EquipeImage/Pandore
3. http ://cimg.sourceforge.net

Deuxième partie

Les pré-traitements

Chapitre 3

La détection du couffin

3.1 Problématique

Dans les vidéos du corpus deux caractéristiques de la prise de vue posent problème pour détecter les mouvements du nouveau-né. Tout d'abord la caméra n'est pas positionnée de manière à offrir une vue de dessus, mais une plongée en diagonale. Du fait de la perspective, l'image du bébé est déformée. De plus, la vidéo ne montre pas que le couffin et l'enfant, mais aussi une partie de la pièce autour de lui. Les objets présents, et les éventuelles personnes passant dans le champs de la caméra, causent des parasites réduisant la fiabilité de la détection.

Afin de résoudre ces deux problèmes en même temps, il faut donc recadrer la vidéo sur le couffin. Cela se fera en deux étapes :

- la détection des bords du couffin
- le recadrage de l'image

3.2 La détection de lignes

3.2.1 Les contours

De nombreuses techniques de traitement d'image permettent de détecter les lignes droites, utilisant par exemple des méthodes de détection de contour (gradient par filtrage ou par masque, passage par zéro du laplacien...). Elles permettent toutefois rarement d'obtenir directement une caractérisation du contour : est-ce une ligne droite ou une courbe ? sa longueur est-elle importante ?

En effet la plupart ne fournissent qu'une image de contour, et il faut alors exécuter d'autres traitements afin d'analyser et de donner un sens à ces countours. L'objet que nous cherchons à détecter peut-être représenté par un trapèze, dont deux des côtés sont plus ou moins horizontaux. Nous voulons donc trouver dans l'image jusqu'à quatre lignes droites, positionnées dans une configuration spécifique.

3.2.2 La transformée de Hough

La transformée de Hough permet de répondre à ces questions [2]. Il s'agit d'une opération permettant de reconnaître les lignes droites dans une image, voire, sous sa forme généralisée, n'importe quelle autre forme. Afin de trouver les alignements, un détecteur de contour est d'abord appliqué à l'image (gradient, sobel, canny...). Chaque point de contour va être représenté par une sinusoïde dans l'espace de Hough, représentant en coordonnées polaire ρ et θ toutes les droites pouvant passer par ce point. Les points d'intersection des sinusoïdes correspondent à une droite reliant les points dans l'image. Plus une intersection est commune à de nombreuses sinusoïdes, plus cette droite est importante. L'espace est discrétisé, et devient un espace accumulateur où chaque point représente une quantité de vote pour une droite. Il faut cependant se souvenir qu'il existe toujours une droite reliant deux points, et que par conséquent l'espace de Hough est saturé par la détection d'absolument toutes les lignes de l'image, y compris les plus insignifiantes. Il faut donc établir un seuil à partir duquel un alignement est considéré comme pertinent et digne d'intérêt.

En partant du postulat que les contours du couffin font partie des lignes principales de l'image, nous allons trier les résultats par importance. D'une manière empirique, il faut prendre au moins les dix premières droites afin d'être certain d'avoir les bords du couffin. Il suffit de réduire ce nombre à un maximum de quatre en utilisant une heuristique basée sur la position et l'angle, par exemple les bords supérieurs et inférieurs sont plus ou moins horizontaux, le bord gauche est toujours dans la partie gauche de l'image...

Chapitre 4

Redressement et découpage

4.1 Les coordonnées spatiales

Une fois la transformée de Hough et la discrimination des bords effectués, nous nous retrouvons en possession des quatre coordonnées polaires des droites. Cependant, le repère de notre image n'est pas en coordonnées polaires. Afin de faciliter les traitements suivants, nous avons besoin des coordonnées des quatres coins du couffin dans le repère de l'image.

Considérons deux droites dA et dB, respectivement de coordonnées polaires (ρ_A, θ_A) et (ρ_B, θ_B), représentant deux bords adjacents du couffin.

Considérons le point $O(x_O, y_O)$ centre de l'image, les points $A(x_A, y_A)$ et $B(x_B, y_B)$ les intersections des droites dA et dB avec leurs hauteurs en O, et le point I l'intersection des deux droites :

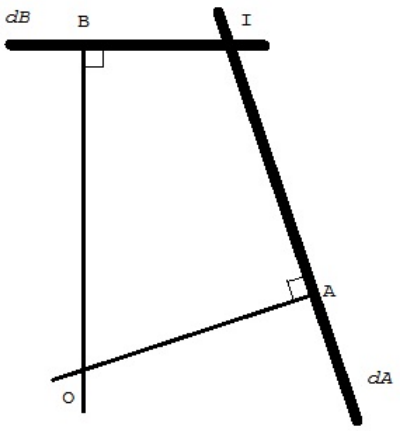

FIGURE 4.1 – Intersection des deux droites

Le but est d'obtenir les coordonnées du point $I(x_I, y_I)$.

Les coordonnées du point A sont :

$$x_A = \rho_A \cos\theta_A + x_O$$
$$y_A = \rho_A \sin\theta_A + y_O$$

et de même pour B.
L'équation de la droite dA est alors de la forme :

$$x + \frac{\sin\theta_A}{\cos\theta_A} y + c = 0$$

Comme A appartient à dA :

$$\rho_A \cos\theta_A + x_O + \frac{\sin\theta_A}{\cos\theta_A}(\rho_A \sin\theta_A + y_O) + c = 0$$

donc

$$c = -\rho_A \cos\theta_A - x_O - \frac{y_O \sin\theta_A + \rho_A \sin^2\theta_A}{\cos\theta_A}$$

L'équation de dA est alors :

$$y = \frac{\cos\theta_A(-x + \rho_A \cos\theta_A + x_O)}{\sin\theta_A} + y_O + \rho_A \sin\theta_A$$

et de même pour dB.
Le point d'intersection I appartient à dA et à dB, d'où l'égalité :

$$\frac{-x_I \cos\theta_A(-x + \rho_A \cos\theta_A + x_O)}{\sin\theta_A} + y_O + \rho_A \sin\theta_A$$
$$=$$
$$\frac{-x_I \cos\theta_B(-x + \rho_B \cos\theta_B + x_O)}{\sin\theta_B} + y_O + \rho_B \sin\theta_B$$

Les y_0 s'annulent. Puis en développant et en réduisant, nous obtenons :

$$x_I = \frac{\rho_B \sin\theta_A - \rho_A \sin\theta_B}{\sin(\theta_A - \theta_B)} + x_O$$

et en réutilisant les équations de dA et dB :

$$y_I = \frac{\cos\theta_A(-x_I + \rho_A \cos\theta_A + x_O)}{\sin\theta_A} + y_O + \rho_A \sin\theta_A$$

et

$$y_I = \frac{\cos\theta_B(-x_I + \rho_B \cos\theta_B + x_O)}{\sin\theta_B} + y_O + \rho_B \sin\theta_B$$

On utilisera la deuxième si $sin\theta_A = 0$.

Nous obtenons ainsi les coordonnées spatiales de chacun des coins du couffin à partir des coordonnées polaires des bords.

4.2 Correction de la perspective

4.2.1 La déformation

Le but est de corriger la déformation causée par l'angle de prise vue de la caméra. Recontruire de manière précise les dimensions d'un objet 3D depuis une image 2D est un problème complexe de reconstruction et d'estimation de pose.

Après des tests manuels sur les vidéos, il est cependant apparu que le ratio de déformation était sensiblement toujours le même. Nous nous contentons donc de la corriger par une simple homographie, c'est-à-dire une transformation projective avec des paramètres prédéterminés.

4.2.2 Implémentation

Nous utilisons les formules issues de l'article de Paul Heckbert [3]. L'avantage d'une telle technique est sa rapidité :

1. détection et calcul des coordonnées des coins du couffin
2. résolution du système linéaire à partir de ces coordonnées (cf. annexe)
3. calcul de la taille de l'image corrigée
4. dans chaque pixel de l'image corrigée, on copie le pixel correspondant de la vidéo en transformant les coordonnées suivant le système linéaire

Seule l'étape 4 est répétée pour chaque image. Il ne s'agit alors que d'un simple calcul mathématique de quelques multiplications et additions, suivi d'une opération de copie.

Chapitre 5

La détection du nouveau-né

5.1 La silhouette

5.1.1 Les vidéos du corpus

Une fois les traitements précédents effectués, nous sommes en possession d'une image rectangulaire ne montrant que le nouveau-né à détecter dans son couffin. Trouver un algorithme permettant de ne récupérer que les contours de celui-ci n'est pas un problème en soit : de nombreuses méthodes existent [4]. Nous utiliserons ici une addition du filtre de sobel à plusieurs opérations de morphologie mathématique.

Cependant, le problème vient de l'aspect fortement disparate des vidéos du corpus : la scène n'est quasiment jamais la même. Tout d'abord, le bébé peut-être habillé différement, allant de la simple couche à la gigoteuse complète. La silhouette peut donc ne plus être qu'un rectangle déformé et surmonté d'un demi cercle. Pour certaines vidéos, détecter l'enfant implique des manipulations de l'espace de couleur, car il est impossible de se baser sur les teintes. Ni les vêtements, ni les draps ne sont fixes. Il arrive parfois que la qualité de la vidéo rende difficile à dicerner le vêtement du drap de même couleur ou de même motif, même pour l'oeil humain.

Par conséquent, si l'on arrive toujours à obtenir une silhouette, celle-ci est rarement assez précise et constante pour permettre une détection fiable de la posture du nouveau-né, et la méthode utilisée est trop fortement liée à une vidéo spécifique pour être utilisée de manière générique.

5.1.2 Les images thermiques

Les images infrarouges résolvent d'elles-mêmes ce problème de segmentation fond/sujet : seul le bébé émet de la chaleur dans l'image. Obtenir une silhouette précise n'est alors qu'un problème de seuillage. Les zones n'étant pas à la température ambiante appartiennent très certainement à l'enfant. De plus, sur une

image thermique les contours des objets sont marqués par une ligne plus froide, et donc plus facilement repérable.

5.2 Squelettisation

La silhouette seule ne permet pas d'estimer la position dans laquelle se trouve le nouveau-né. Afin de faciliter cette tâche, nous tentons donc de définir la position du squelette. Diverses techniques de squelettisation existent : simulation du front enflammé, morphologie mathématique, calcul analytique.

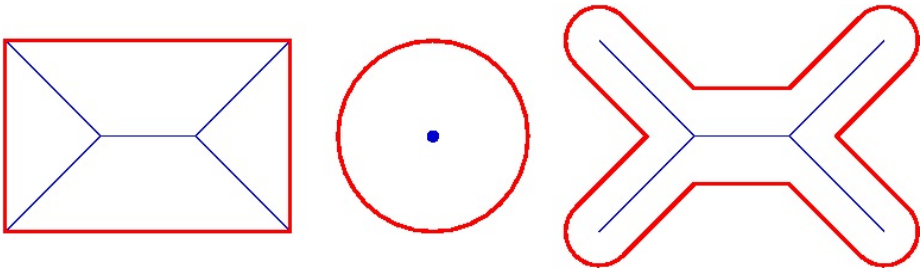

FIGURE 5.1 – Exemples de squelettes de forme simples

La méthode utilisée permet d'obtenir le squelette tout en conservant la silhouette : il s'agit du calcul de la carte de distances. Elle consiste à associer à chaque point de la silhouette la distance au point du contour le plus proche. Les maxima locaux correspondent aux points du squelette (cf. annexe).

Troisième partie

La position et le mouvement

Chapitre 6

Modelisation du corps humain

6.1 Les modèles

6.1.1 Modèles courant

De nombreuses modèlisations du corps humain existent dans la littérature [5] [6], allant de la simple chaîne cinématique représentant le squelette, à l'ensemble complexe de volumes paramétriques constituant chaque muscle. Le choix d'une modélisation provient souvent du mode de capture de mouvement utilisé et de l'application visée : un exosquelette animera plus facilement une modèle squelettique, alors qu'un système à marqueurs optiques correspondra mieux à un modèle volumique.

FIGURE 6.1 – Capture de mouvement par exosquelette (Gipsy), et par marqueurs optiques (Vicon)

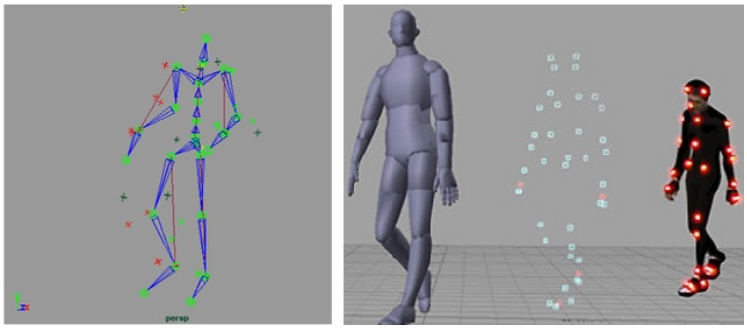

FIGURE 6.2 – Modèles de chaîne cinématique et de volume complexe

6.1.2 Modèle adopté

Ces modèlisations complexes sont lourdes à manipuler, et peu adaptées pour une application simple et expérimentale telle que celle de ce projet. La chaîne cinématique simple représentant le squelette suffit dans notre cas, puisque nous cherchons à suivre les mouvements des membres et à déterminer les températures correspondantes.

Le modèle construit est un ensemble de 15 points représentant les épaules, les hanches, les coudes, les poignets, les genoux, les pieds, le pelvis, le cou, et le centre de la tête. Ces points sont reliés deux à deux par un segment afin d'avoir une représentation visuelle cohérente. Ils possèdent chacun un ensemble de coordonnées spatio-temporelles $\{x, y, z, t\}$, permettant ainsi de représenter les mouvements.

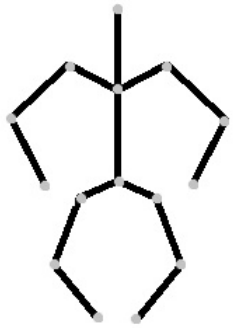

FIGURE 6.3 – Chaîne cinématique simple

6.2 Les contraintes

Le squelette construit doit possèder un certain nombre de contraintes physiques. Certaines sont liées à la physiologie du corps humain : angle maximal d'une articulation, rigidité des os. D'autres viennent de la spécificité de l'application et du sujet : le nouveau-né est toujours allongé, les bébés adoptent naturellement une certaine position, de par leur développement, mais aussi en partie à cause des vêtements et de la rigidité de la couche.

Cependant, ces dernières ne sont pas aussi strictes. Afin de reflèter cette souplesse, les contraintes du modèle ne sont donc pas implémentées par des conditions strictes, par exemple angle maximal du coude inférieur à 180 degrès. Nous introduisons ici la notion d'énérgie, qui sera capitale dans le chapitre suivant.

Chaque contrainte sur le squelette se voit attribuer une zone de confort. Il peut s'agir d'une distance, d'un angle, d'une orientation, maximale ou minimale. Tant que la valeur courante reste dans cette zone, l'énergie reste faible. S'approcher des frontières la fait augmenter, et les dépasser rend sa croissance exponentielle.

Chapitre 7

Détection de posture par squelette actif

7.1 Principe de fonctionnement

7.1.1 Snake

Une méthode de détection couramment utilisée pour trouver les contours d'un objet est celle du contour actif, ou *snake* [7]. Une série de points mobiles formant une courbe sont définis. Cette courbe se déplace à chaque itération jusqu'à épouser la forme de l'objet. Cette dynamique est basée sur la notion d'énergie interne et externe, le but étant de minimiser l'énergie totale présente le long de la courbe. Des contraintes permettent de conserver une courbe lisse avec des points équidistants tout en laissant un certain champ libre pour les déformations. L'énergie interne correspond à la morphologie et aux caractéristiques de la courbe (courbure, longueur...). L'énergie externe provient de l'image, les critères sont variables (présence de bords marqués, bruit...).

FIGURE 7.1 – Détection de contour par snake

7.1.2 Adaptation au problème

Comme nous pouvons le remarquer, notre squelette possède certains points communs avec le snake des contours actifs : il est formé de points reliés entre eux, et nous cherchons à lui faire prendre une certaine forme. Son énergie interne peut alors être définie comme la somme des énergies des contraintes :

$$E_{int} = \sum_{c \in C} E_c$$

où C est l'ensemble des contraintes des articulations du squelette.

L'énergie externe est liée à l'image. Sur notre carte des distances, le squelette de la silhouette correpond en partie à la projection 2D du squelette réel. Cette énergie correspond donc à la manière dont le projeté de notre squelette virtuel se calque sur l'image. Nous pouvons donc définir l'énergie du segment entre deux articulations :

$$E_s = \frac{l}{\sum_{k=0}^{l} I_k}$$

où l est la longeur du segment projeté sur l'image et I_k la valeur sur la carte des distances à un pixel appartenant au segment. Sur la carte, les valeurs les plus fortes correspondent au squelette de la silhouette, et par conséquent l'énergie d'un segment est minimale quand il est superposé à ces pixels.

Ainsi, l'énergie externe devient :

$$E_{ext} = \sum_{s \in S} E_s$$

où S est l'ensemble des segments du squelette.

7.2 Implémentation

7.2.1 La minimisation de l'énergie

Il s'agit donc de minimiser l'énergie globale :

$$E = \alpha E_{int} + (1 - \alpha) E_{ext}$$

où α est un paramètre de réglage permettant de priviligier une énergie sur l'autre ($0 \leq \alpha \leq 1$).

Après l'initialisation du squelette virtuel à une position probable, basée sur la posture généralement adoptée par un nouveau-né et adaptée pour correspondre le plus possible à la silhouette, l'implémentation suit le schéma suivant :

1. calcul de l'énergie
2. léger déplacement de chaque point du squelette dans l'espace, avec pour chaque le calcul du changement énergétique induit.
3. déplacement du squelette dans la position où l'énergie est minimale.

Les étapes 2 et 3 sont alors répétées jusqu'à stabilisation de l'énergie.

7.2.2 Problèmes

Il existe de nombreuses implémentations optimisées du *snake*, et les adapter ne présente pas de difficultés majeures. Cependant, pour les deux méthodes, le problème principal posé vient de l'initialisation des points : il peut arriver d'obtenir un minimum local de l'énergie qui n'est pas la solution souhaitée, tout simplement car la position de départ ne permettait pas d'y arriver.

De plus, même si la solution optimale est trouvée, c'est-à-dire que la vraie position de l'enfant est détectée, exécuter la minimisation sur chaque image de la vidéo ne donne pas un mouvement précis et fluide : en effet la déformation de la silhouette n'est pas régulière et par exemple un amincissement peut produire un déplacement du squelette virtuel dans le sens opposé au mouvement réél.

Chapitre 8

Initialisation et précision

8.1 Semi-supervisée

Une solution au problème d'initialisation consiste à permettre à l'utilisateur de relancer le processus du squelette actif en cas d'échec, en le faisant partir d'une autre position. Celle-ci peut être générée aléatoirement à partir d'une graine, mais il serait aussi possible de guider la détection manuellement.

Ces solutions, si elles permettent une meilleure détection, ne doivent être utilisée qu'en dernier recours, car elles restent peu pratiques. Nous devons donc nous pencher sur une autre approche, exploitant la quatrième dimension disponible : le temps.

8.2 Le flot optique

Le calcul du flot optique consiste à estimer le mouvement apparent dans une vidéo en trouvant les pixels correspondant dans les images consécutives, avec l'hypothèse que l'intensité (ou la couleur) est conservée au cours du déplacement. Ce résultat peut ensuite être utilisé dans de nombreuses applications, telles que la segmentation, la poursuite, l'interpolation temporelle...

De nombreuse méthodes de calcul du flot optique ont été proposées depuis 1981 par Horn et Schunk [8], Lukas et Kanade, Nagel et Enkelmann et de nombreux autres depuis. Elles se basent sur différents opérateurs de régularisation, des analyses multi-échelles, la variation totale, la corrélation.... Il s'agit d'un calcul lourd, mais des méthodes permettent de l'obtenir en temps réel [9].

Dans le cadre de notre application, le but recherché n'est pas d'obtenir une estimation optimale en temps-réelle, mais de savoir si l"utilisation du flot optique permet d'améliorer la détection de la position du nourrisson. Nous nous contenterons par conséquent d'une implémentation de type Horn et Schunk.

8.3 Utilisation de l'estimation

8.3.1 Initialisation

Le flot optique permet une meilleure initialisation du squelette actif qu'une image fixe. En effet, certains problèmes liés à l'occlusion ou à la superposition peuvent être levés grâce au mouvement. Par exemple, si le bras est sur le ventre, avec un vêtement empéchant de le distinguer (couleur, motif...), la nature du mouvement permet de l'identifier plus facilement, et de positionner les points d'articulation.

8.3.2 Optimisation de la convergence

Sans l'utilisation du flot optique, les articulations sont déplacées dans toutes les directions afin de trouver l'énergie minimale. L'estimation du mouvement permet de ne rechercher une énergie minimale que dans un nombre réduit de directions. Plus le vecteur de déplacement est fort, moins nous allons chercher dans les directions opposées. Nous avons donc un gain en temps de calcul proportionnel au mouvement.

Par ailleurs, les mouvements du squelette sont moins soumis à des aberrations, comme des retours saccadés en arrière dus à la déformation de la silhouette. En effet, la régularisation du flot optique entraîne un lissage du mouvement puisque l'énergie globale dépend à présent aussi du flot optique :

$$E = \alpha E_{int} + \beta E_{ext} + \gamma E_{fo}$$

où α, β et γ sont des paramètres de réglage permettant de priviligier une énergie sur les autres, et E_{fo} est l'énergie du flot optique, issue de la soustraction des vecteurs de déplacement du point du squelette et de mouvement du pixel.

Quatrième partie

Applications

Chapitre 9

Le mouvement dans les vidéos du corpus

9.1 Les contraintes et les problèmes

9.1.1 But du corpus

Le corpus présente la fois des enfants sains et des enfants atteints d'une maladie. Après avoir effectué une détection et une mesure des mouvements sur chacune des vidéos, il serait intéressant de pouvoir déterminer si les caractéristiques des gestes relevés peuvent permettre de diagnostiquer le problème.

De nombreuses méthodes d'apprentissage, liées au domaine de l'intelligence artificielle et à la fouille de données, sont capables de rechercher et d'extraire les points communs et les différences entre les vidéos, et d'en déduire une classification.

9.1.2 Précision et fiabilité de la détection

De telles méthodes demandent un grand nombre de données afin d'obtenir un résultat fiable. Cela implique donc de réaliser la détection des mouvements sur une large quantité de vidéos, et de manière automatisée. Cependant la nature disparate des vidéos du corpus ne permet pas une détection suffisament fiable et précise en utilisant la méthode du squelette actif en l'état actuel.

Nous disposons cependant tout de même d'une information : la zone dans laquelle nous devons analyser les mouvements est définie par le couffin.

9.2 Calcul du mouvement et extension à la classification

En reportant notre attention sur cette zone, nous pouvons nous demander si une mesure précise du mouvement de chacun des membres du nouveau-né est nécessaire dans un but de classification.

9.2.1 Mouvement global

En effet, nous avons vu que le flot optique permet d'obtenir une estimation correcte du mouvement dans la vidéo. Cette estimation contient déjà toutes les données qui pourraient être extraite par la détection par squelette actif. Ces données ne possèdent juste aucune notion de sémantique : nous ne savons pas si ce qui bouge est un bras ou une jambe, mais nous savons comment il bouge. Selon M. Jouen, la nature du mouvement est souvent ce qui caractérise le diagnostique : trop saccadé, trop fixe... Il serait donc peut-être possible de n'utiliser que le mouvement global.

9.2.2 Découpage en zone

Par ailleurs si le mouvement global ne suffit pas, le découpage du couffin en zone pourrait aider à préciser les données. En effet, le bébé est généralement placé au centre du couffin. Ainsi, séparer l'image en quatre zones permettrait de connaître les mouvements globaux d'une jambe précise. Il semble aussi arriver qu'un enfant malade vienne à se recroqueviller dans un coin du lit. L'absence de mouvement dans les trois quarts de l'image pourrait donc aussi être une donnée utilisable pour discriminer les vidéos.

Chapitre 10

Position et température

10.1 Les vidéos thermiques

La caméra thermique venant d'être acquise, aucune donnée réelle n'a encore pu être exploitée. Cependant, certaines applications et tests restent possible avec des images et des vidéos artificielles.

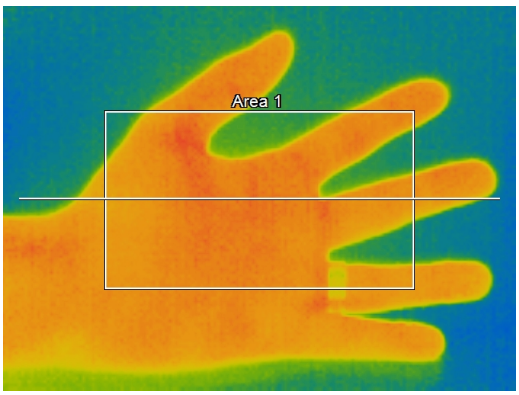

FIGURE 10.1 – Exemple d'image thermique issue de la caméra Optris PI

Bien évidemment, les images de ce type, même si elles sont pratiques pour visualiser le résultat, sont en fausses couleurs qui dépendent de la palette choisie et de la configuration de la caméra. Le travail ne sera donc pas effectuée sur celle-ci, mais sur les données réelles. L'image est la même, mais les pixels, au lieu de contenir une couleur, contiennent une température.

10.2 Températures

10.2.1 Relevés globaux

À partir des données thermiques, plusieurs valeurs intéressantes sont à relever :

1. la température minimale, que l'on peut supposer appartenir au couffin.
2. la température maximale, que l'on peut supposer appartenir au bébé.
3. la température moyenne pouvant aider à déterminer le seuil de chaleur correspondant à l'enfant.

Évidemment, ces suppositions dépendent de l'environnement. Si elles peuvent s'avérer exacte pour un nourrisson dans un couffin, elles ne les sont plus forcément pour un nouveau-né dans une couveuse dont la température ambiante peut atteindre les 37 ° C.

Les données seules ne suffisent donc pas à effectuer des relevés. L'estimation de certaines valeurs spécifiques demandent des mesures extérieures, comme par exemple les réglages de la couveuse et de la caméra thermique.

10.2.2 Relevés locaux

Nous pouvons également utiliser notre détection de posture de l'enfant pour établir des relevés de températures locaux. La zone sur laquelle sont effectués les calculs (maximale, minimale, moyenne...) est alors déterminée par les frontières de la silhouette autour du segment du squelette, auquelles s'ajoutent la bissectrice de l'articulation.

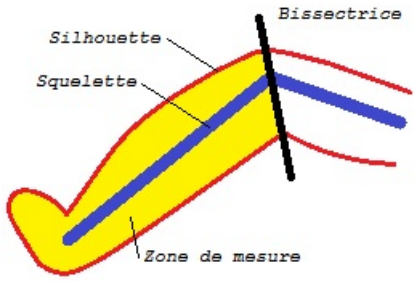

FIGURE 10.2 – Zone de mesure locale

Il faut bien entendu prendre en compte le problème de l'occlusion et de la superposition. Nous considérerons donc en premier les membres les plus proches de la caméra. Ainsi quand un pixel de température lui a été attribué, il ne sera pas attribué à une partie cachée.

10.3 Utilisation des relevés

Tout d'abord, des courbes de températures pour chaque relevé local ou global peuvent être établies. Ces courbes permettent d'avoir un historique. De plus les variations de températures peuvent être liées à certains diagnostiques [1].

Il est aussi possible de mettre un place un système de surveillance. Ainsi, les médecins peuvent être informés si la température chute de manière dangereuse. Cette application est notamment intéressante dans le sens où l'alarme peut être déclenchée en cas de perte de température dans un seul membre, ce qui ne serait pas le cas avec une sonde unique placée sur le corps de l'enfant.

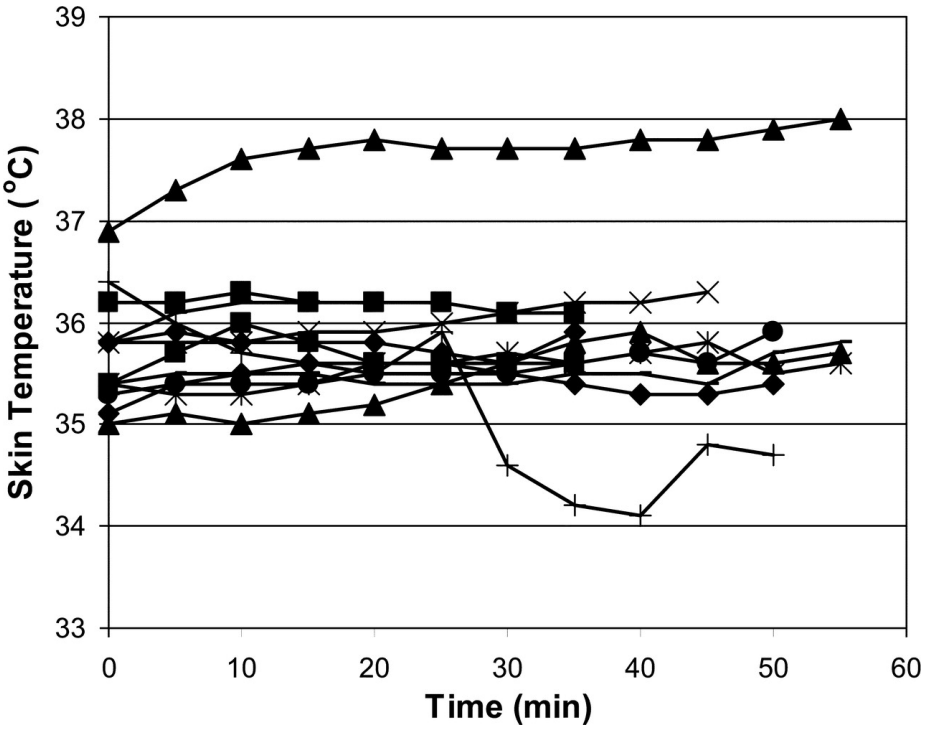

FIGURE 10.3 – Relevé de températures

Conclusion

Les résultats

Détection du mouvement

La méthode de détection de la postion du nouveau-né se rélève assez efficace, mais elle dépend trop du pré-traitement de détection de la silhouette. En effet, si un traitement unique n'a pas été trouvée pour les vidéos du corpus, les vidéos thermiques pourraient apporter la solution au problème. Dans le cas où les contours sont précis, le squelette peut prendre rapidement la bonne position. Le programme C++, dont le code n'est que très peu voire pas du tout optimisé, parvient à un résultat en quelques secondes (compilation O2, core i5-2500k@3,3GHz). La synchronisation des mouvements se fait alors en temps réel grâce au flot optique pré-calculé.

Il pourrait cependant être possible d'exploiter le mouvement dans les vidéos du corpus. Le flot optique donne en effet une bonne estimation de l'ensemble des mouvements présents.

Exploitation des relevés

L'exploitation des relevés de températures se révèle très efficace pour deux raisons. Tout d'abord, la détection des contours est simplifiée. En effet, une ligne de températures plus froides est présente autour de chaque être vivant, même en cas de superposition, ce qui offre une silhouette naturelle bien plus exploitable que celle basée sur la couleur. De plus, le traitement des données s'avère très rapide. En effet, une fois la délimitation des zones correspondantes aux segments du squelette effectuée, il est possible de calculer la plupart des valeurs intéressantes (minimale, maximale, moyenne...) en un seul parcours de l'image.

Extensions et possibilités

Plusieurs développements et tests s'inscrivent dans la continuité de ce stage pour le projet PréTherm. Dès que les vidéos thermiques seront disponibles, le

développement et l'optimisation des techniques proposées pourront être poursuivis. Les essais avec les données réelles permettront notamment de déterminer certains paramètres, comme par exemple la fréquence des traitements pouvant être atteinte.

De plus, une autre méthode de calcul du flot optique doit être utilisée. En effet, l'algorithme basé sur Horn et Schunk implémenté pour ces tests reste primaire, et d'autres méthodes peuvent offrir de meilleurs résultats. Le coût de calcul de cette méthode doit aussi être étudié, car une utilisation en temps réelle sur les images thermiques (par exemple pour la surveillance des seuils de température) ne pourra pas utiliser un flot optique pré-calculé comme c'est ici le cas, mais devra l'estimer au fur et à mesure.

Enfin, concernant les vidéos du corpus et la classification automatique des diagnostiques, différents essais pourront être réalisés sur un plus grand nombre de vidéos. Il faudra notamment choisir une méthode d'estimation du mouvement, ainsi qu'un algorithme classifieur. Cette tâche représente cependant une grande quantité de travail.

Ressenti du stage

Ce stage a été pour moi l'occasion de découvrir le monde de la recherche. J'ai pu constater ainsi ses avantages, tels que travailler sur des nouvelles techniques à la fois du domaine médical et du traitement d'image, mais aussi ses aléas : les contre-temps imprévus, les avancées à tatons...

Je tiens par ailleurs à remercier M. Chahir, qui a su me guider dans mes recherches tout en me permettant de travailler en autonomie et de faire mes propres constatations.

Cinquième partie

Annexes

Annexe A

Les difficultés des vidéos du corpus

FIGURE A.1 – Ces enfants ne sont pas des nourrissons et ne peuvent pas bouger librement dans le couffin.

FIGURE A.2 – Le drap peut fausser la détection.

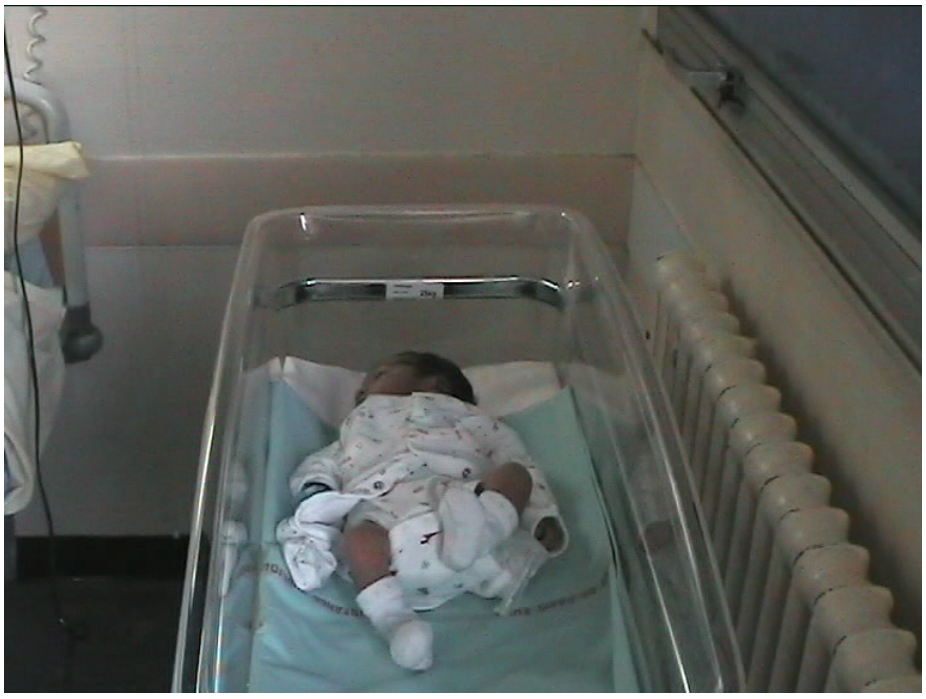

FIGURE A.3 – L'angle de prise de vue rend difficile la détection et le vêtement crée des faux mouvements.

FIGURE A.4 – Le vêtement dissimule les mouvements.

Annexe B

Les images thermiques

B.1 Les bases de la thermographie infrarouge

Suivant la température chaque objet émet une certaine quatité de rayonnement infrarouge. Un changement de température de l'objet est accompagné par une variation de l'intensité du rayonnement.

En 1800, l'astronome William Hersche (1738 - 1822) découvrit par chance les rayons infrarouges. En cherchant à mesurer la chaleur de chaque couleur en projetant la lumière décomposée par un prisme sur une thermomètre à mercure, il s'aperçut que la chaleur la plus forte était du côté du rouge, et que la température augmentait encore après avoir dépassé la zone de lumière visible.

Les thermomètres infrarouge sont des capteurs opto-électroniques. Ils calculent la température de surface d'un objet en se basant sur le rayonnement émis. Le principal intérêt de ce type de capteur est qu'ils permettent une mesure sans contact physique. Ils ont généralement pour base une lentille, un filtre spécial, un capteur, et une partie électronique chargée de l'amplification, de la linéarisation, et du traitement du signal.

B.2 La caméra Optris PI

Spécifications générales du constructeur :
- Environmental rating IP65 (NEMA-4)
- Ambient temperature 0...50 ° C
- Storage temperature -20...70 ° C
- Relative humidity 10...95 %, non condensing
- Material (housing) aluminium, anodized
- Dimensions 45 mm x 45 mm x 65 mm
- Weight 250 g (incl. lens and 1 m USB cable)
- Cable length (USB 2.0) 1 m (Standard), 5 m, 10 m, 20 m
- Vibration IEC 68-2-6 : 3G, 11 – 200Hz, any axis
- Shock IEC 68-2-27 : 50G, 11ms, any axis

Electrical Specifications :
- Power Supply 5 VDC (powered via USB 2.0 interface)
- Current draw max. 500 mA
- Output Process Interface (PIF out) 0-10 V (TObj, TInt,Flag status or Alarm status)
- Input Process Interface (PIF in) 0-10 V (Emissivity, Ambient temperature, Reference temperature,
- Flag control, triggered video or triggered snapshots)
- Digital Input Process Interface Flag control, triggered video or triggered snapshots)
- Digital interface USB 2.0

Measurement Specifications :
- Temperature ranges -20...100 ˚ C
- 0...250 ˚ C
- 120...900 ˚ C
- Spectral range 7.5...13 µm
- Lenses 23 ˚ x 17 ˚ FOV / 6 ˚ x 5 ˚ FOV / 48 ˚ x 37 ˚ FOV
- System accuracy 1) ±2 ˚ C or ±2 %
- Temperature resolution (NETD) 0.08 K with 23 ˚ FOV
- Frame rate 120 Hz
- Warm-up time 10 min
- Emissivity 0,100. . . 1,000 (adjustable via software)

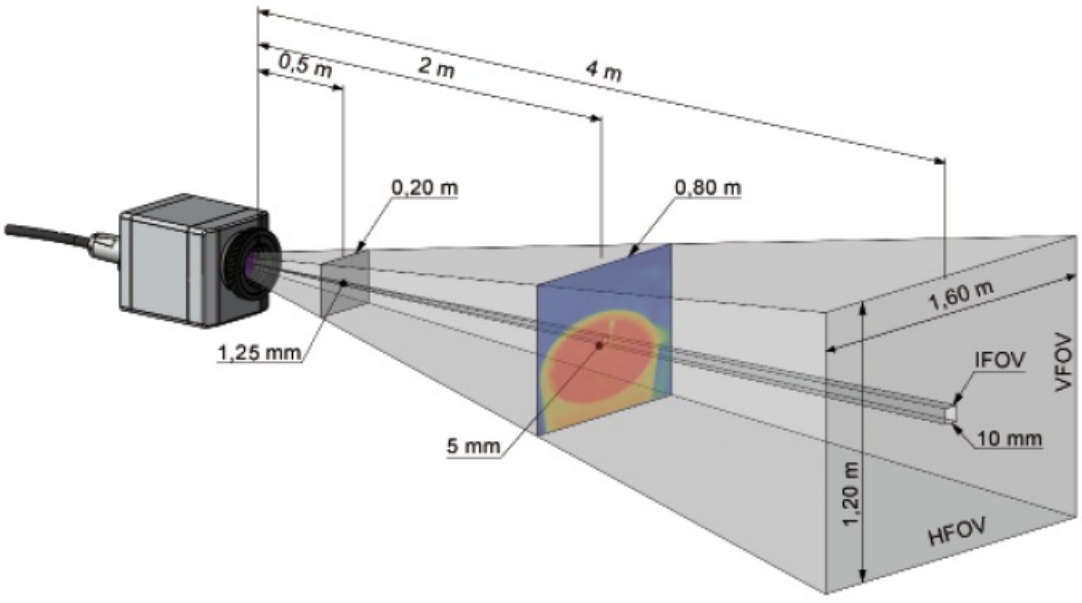

FIGURE B.1 – Champ de mesure avec la lentille standard 23 ˚ x 17 ˚

Annexe C

Le progamme développé pour les tests

C.1 Transformée de Hough

```
1  CImg<> vote(this->width,this->height,1,1,0);
2  CImgList<> grad = img.get_gradient();
3  cimglist_for(grad,l) grad[l].blur((float)alpha);
4  vote.fill(0);
5  cimg_forXY(img,x,y) {
6          const double
7                  X = (double)x-img.width()/2,
8                  Y = (double)y-img.height()/2,
9                  gx = grad[0](x,y),
10                 gy = grad[1](x,y);
11         double
12                 theta = std::atan2(gy,gx),
13                 rho   = std::sqrt(X*X+Y*Y)*std::cos(std::atan2(Y,X)
14                                                        -theta);
15         if (rho<0) { rho=-rho; theta+=cimg::PI; }
16         theta = cimg::mod(theta,thetamax);
17         vote((int)(theta*vote.width()/thetamax),
18                         (int)(rho*vote.height()/rhomax))
19                 +=(float)std::sqrt(gx*gx+gy*gy);
20 }
```

C.2 Résolution du système linéaire pour le recadrage

```
1  vector<double> Redress::getSystem(point<int> P0, point<int> P1,
2                   point<int> P2, point<int> P3) {
3          vector<double> system(8);
4          double sx = (P0.x-P1.x)+(P2.x-P3.x);
```

```
5           double sy = (P0.y-P1.y)+(P2.y-P3.y);
6           double dx1 = P1.x-P2.x;
7           double dx2 = P3.x-P2.x;
8           double dy1 = P1.y-P2.y;
9           double dy2 = P3.y-P2.y;
10
11          double z = (dx1*dy2)-(dy1*dx2);
12          double g = ((sx*dy2)-(sy*dx2))/z;
13          double h = ((sy*dx1)-(sx*dy1))/z;
14
15          system[0]=P1.x-P0.x+g*P1.x;
16          system[1]=P3.x-P0.x+h*P3.x;
17          system[2]=P0.x;
18          system[3]=P1.y-P0.y+g*P1.y;
19          system[4]=P3.y-P0.y+h*P3.y;
20          system[5]=P0.y;
21          system[6]=g;
22          system[7]=h;
23
24          return system;
25  }
```

C.3 Algorithme squelette actif

Version simple avec voisinnage 15 et prise en compte du flot optique :

```
1  float p = getPixelSize(); //pas de voisinage
2  float dx[15] = {0, p, 0, 0,-p, 0, 0, p, p,-p,-p,-p,-p, p, p};
3  float dy[15] = {0, 0, p, 0, 0,-p, 0, p, p, p, p,-p,-p,-p,-p};
4  float dz[15] = {0, 0, 0, p, 0, 0,-p, p,-p, p,-p, p,-p, p,-p};
5  list<point3d<float>* >::iterator it;
6  //pour chaque articulation du squelette
7  for(it=skel.begin() ; it!= skel.end(); it++){
8          point3d<float> * a = (*it);
9          int i = 0, bestI = i;
10         float bestE = energy();//meilleure energie
11         for (i=1; i<15; i++){//pour chaque voisin
12                 //flot optique
13                 if (!directionFO(dx[i], dx[y], dx[z])) continue;
14                 a->x += dx[i];
15                 a->y += dy[i];
16                 a->z += dz[i];
17                 float eTmp = energy();
18                 if (eTmp<bestE){
19                         bestE = eTmp;
20                         bestI = i;
21                 }
22                 a->x -= dx[i];
23                 a->y -= dy[i];
24                 a->z -= dz[i];
25         }
26         a->x += dx[bestI];
27         a->y += dy[bestI];
28         a->z += dz[bestI];
29 }
```

Annexe D

Exemples de résultats

FIGURE D.1 – Transformée de Hough

FIGURE D.2 – Recadrage de l'image

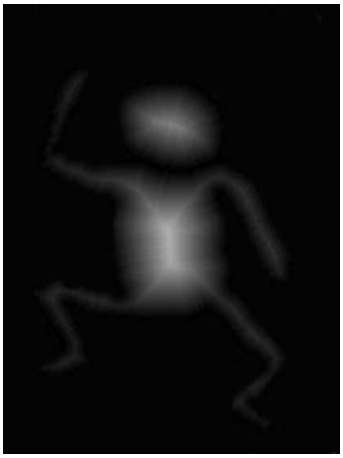

FIGURE D.3 – Squelettisation par carte de distances

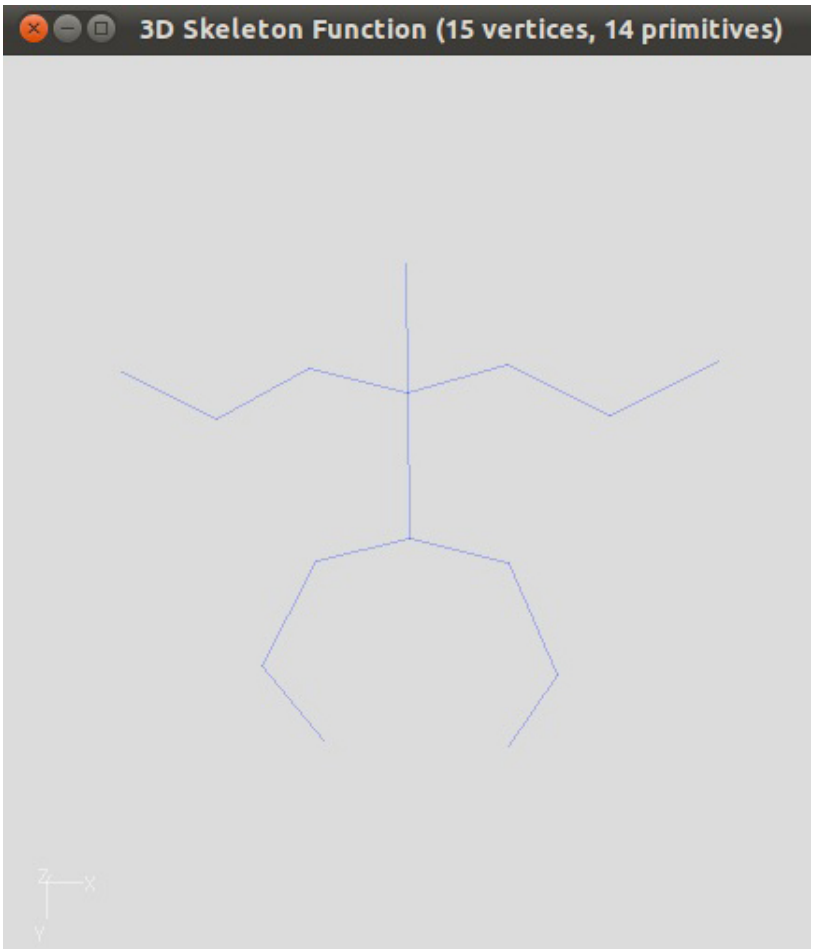

FIGURE D.4 – Modèle du squelette actif affiché par CImg (posture par défaut non initialisée)

Bibliographie

[1] Caroline L. Hollingsworth MD Elizabeth Bradsher RN Melissa E. Danko MD Stephanie M. Crosby LPN Ronald N. Goldberg MD David T. Tanaka MD Henry E. Rice, MD and RNC NNP Robin B. Knobel, PhD. Infrared thermal imaging (thermography) of the abdomen in extremely low birthweight infants. *Journal of Surgical Radiology*, July 2010.

[2] J. Illingworth and J. Kittler. A survey of the hough transform. *Computer Vision, Graphics, and Image Processing*, 44 :87–116, October 1988.

[3] Paul Heckbert. Projective mappings for image warping. *15-869, Image-Based Modeling and Rendering*, September 1999.

[4] Djemel Ziou and Salvatore Tabbone. Edge detection techniques - an overview. *International Journal of Pattern Recognition and Image Analysis*, 8 :537–559, 1998.

[5] David Knossow. *Analyse et capture multi-caméra du mouvement humain.* PhD thesis, Institut National Polytechnique de Grenoble, 2007.

[6] Jungong Han Weilun Lao and Peter H.N. de With. Fast detection and modeling of human-body parts from monocular video. In *AMDO*, September 2008.

[7] ANDREW WITKIN MICHAEL KASS and DEMETRI TERZOPOULOS. Snakes : Active contour models. *International Journal of Computer Vision, 321-331*, 1988.

[8] Berthold K.P. Horn and Brian G. Rhunck. Determining optical flow. *Artificial Intelligence.*

[9] Kohlberger T. Schnorr C. Bruhn A., Weickert J. ; Feddern C. Variational optical flow computation in real time. *IEEE Transactions on Image Processing*, 14 :608 – 615, May 2005.

Résumé :

Ce rapport présente le déroulement et la conclusion du travail effectué par Florian Jacomme lors du stage de validation du Master 2 Recherche Langage, Image et Document, effectué à l'Université de Caen Basse Normandie sous la tutelle de M. Chahir Youssef. Le sujet de ce stage portait sur le traitement et l'analyse de vidéo, d'une part d'un point de vue détection des mouvements, et d'autre part de l'utilisation d'images thermiques. Ce travail est effectué dans le cadre du projet PréTherm, visant à permettre l'analyse et la détection de la douleur chez les nouveau-nés prématurés par une analyse des images thermiques. Durant ce stage ont été effectués des travaux de prétraitements visant à détecter le couffin (transformée de hough) et à recadrer l'image sur celui-ci, une détection de la posture du nourrison par un modèle de squelette actif, une étude de la corrélation entre cette détection et les relevés de températures issus des images thermiques et une étude de l'utilisation de la détection de mouvement dans le cadre d'une classification d'un corpus de vidéos médicales.

Mots clés : analyse, traitement, recherche, vidéo, médical, PréTherm, mouvement, infrarouge, thermique, flot optique, squelette actif